

# A Variational Bayes Approach to Decoding in a Phase-Uncertain Digital Receiver

Arijit Das<sup>†</sup> and Anthony Quinn<sup>\*</sup>

*Department of Electronic and Electrical Engineering  
Trinity College Dublin*

E-mail: <sup>†</sup>dasa@tcd.ie

<sup>\*</sup>aquinn@tcd.ie

---

*Abstract* — This paper presents a Bayesian approach to symbol and phase inference in a phase-unsynchronized digital receiver. It primarily extends [10] to the multi-symbol case, using the variational Bayes (VB) approximation to deal with the combinatorial complexity of the phase inference in this case. The work provides a fully Bayesian extension of the EM-based framework underlying current turbo-synchronization methods, since it induces a von Mises prior on the time-invariant phase parameter. As a result, we achieve tractable iterative algorithms with improved robustness in low SNR regimes, compared to the current EM-based approaches. As a corollary to our analysis we also discover the importance of prior regularization in elegantly tackling the significant problem of phase ambiguity.

*Keywords* — Variational Bayes approximation, phase synchronization, phase ambiguity resolution, von Mises distribution, soft decoding

---

## I INTRODUCTION

Communication Systems face uncertainties of the transmitter and receiver oscillators, the noisy dispersive nature of the channel, among others for transferring messages from the transmitter to the receiver [11]. Incorrect alignment of the transmitter and receiver oscillators shows itself as a rotation of the received data on the complex plane and renders decoding impossible. The process of aligning these oscillators is known as synchronization and is an important aspect in the process of decoding [2].

Current state of the art in dealing with this problem is a joint iterative estimation of the phase and the symbol sequence, in an algorithm known as turbo-synchronization [3]. This algorithm is based on the framework of an EM-algorithm where probability distribution on the symbol decisions are used recursively to improve the estimate of phase, and vice-versa. This use of probability distributions over the possible symbols while making a decision, defined as a "soft" decision, was the novelty in this approach [3]. The algorithm could also be defined under other frameworks, such as

one based on the sum-product algorithms [4], [7].

The EM-algorithm however, considers certainty-equivalence measures on the parameter values which lead to the loss information about their statistical uncertainty. This work aims to generalize this framework into a fully Bayesian treatment, which is flexible enough to "learn" the varied channel conditions along with the uncertainties associated with the modelled parameters. This was however partially achieved in [9], wherein the authors used a uniform prior over the phase. Good Bayesian learning models take into consideration all the statistical information, the data has to offer and therefore advice better decisions under adverse conditions.

This paper presents the advantage of using a generalization of the uniform distribution on the unit circle, the Von-Mises distribution in resolving phase ambiguity. Literature consists of either using a huge number of pilot symbols or modified likelihood estimators [4] which are generally inefficient with respect to power and time, to deal with the problem. In the course of our analysis, we find a natural elegant solution to the problem, which does not add any computational burden to the syn-

chronization routine and exalts the importance of prior regularization in Bayesian Analysis.

As constellation alphabets belong to finite fields, the receiver data models are inevitably mixture distributions which are not in the exponential family. Finite dimensional sufficient statistics [5] and conjugate prior distributions [1] are a unique characteristic to the exponential family. They are important in devising recursive learning algorithms. Unfortunately as this is not the case, it leads to models with substantially expensive inference procedures. A deterministic approximation method, known as Variational Bayesian (VB) Method [14], is applied in this paper to deal with the problem.

The rest of the paper is structured as follows. Section II defines the system model encompassing unknown phase and symbol sequence. Section III describes the single symbol transmission inference. Section IV describes the combinatorial complexity, the Variational Bayes Method and how it is applied to the Multi-Symbol transmission inference. In section V, numerical results for the performance of the described receiver and its comparisons with the current state of the art are provided. Section VI discusses the results and its implications. Section VII presents the concluding remarks and future course of actions.

## II SYSTEM MODEL

Structured redundancy is usually added in the transmitted data [2], in a process known as coding, to protect from the noisy nature AWGN channel. Equal partitions of the binary message stream formed from uniformly sampled and quantized analog signals, are bijectively mapped to a complex symbol  $a = [a_1, \dots, a_m, \dots, a_M]'$  from a constellation alphabet  $\mathcal{A}$  of size  $M$ . These symbols are then modulated onto a carrier signal which can be sent through a physical channel. On the receiver side, the reverse operation called demodulation attempts to decode or identify the transmitted symbol sequence, which are now corrupted by noise [3].

We observe the data vector  $x_i \in \mathbb{C}^n$  during every symbol period with  $i \in 1, \dots, K$  with  $\mathbf{x}_i = [x'_1, x'_2, \dots, x'_i]'$  and assume the prior probabilities corresponding to the symbol vector  $a$  to be  $\alpha = [\alpha_1, \dots, \alpha_M]'$ .  $\phi$  represents the phase difference of the received signal to the transmitted, with noise  $r$  in an AWGN channel [10]. Ease of inferential analysis dictates the augmentation of the model using a latent variable. Let it be  $l_i = [l_{1,i}, \dots, l_{M,i}]' \in \{\epsilon_M(1), \dots, \epsilon_M(M)\}$ , where  $\epsilon_M(m) = [\delta(m-1), \dots, \delta(m-M)]'$  and  $\delta$  is the Kronecker Delta. Hence  $l_i$  is a vector pointer to the component active at the  $i^{th}$  symbol period. The data generation model for such a symbol transfer

in a single symbol period can be defined as:

$$f(x_i|\phi, r, \alpha) = \sum_{l_i} f(l_i|\alpha) f(x_i|a, \phi, r, l_i) \quad (1)$$

$$f(x_i|\phi, r, l_i) \equiv \mathcal{CN}_x(l_i a e^{j\phi}, r) \quad (2)$$

$$f(l_i|\alpha) \equiv \mathcal{Mu}_{l_i}(1, \alpha) = l_i' \alpha \quad (3)$$

where  $\mathcal{CN}$ : Complex Normal and  $\mathcal{Mu}$ : Multinomial distributions,  $\phi \in [-\pi, \pi]$ ,  $r \in (0, \infty)$ ,  $g_i = [s_1 e^{j\omega}, \dots, s_n e^{j\omega n}]'$  with  $s_i \in \{1, 2, \dots, n\}$  as the pulse sequence and  $\omega$  the carrier digital frequency which are assumed to be known. In this analysis we assume known SNR, based on which we can calculate the noise variance,  $r$ .

### a) Prior Distribution to Phase: Von-Mises Distribution

As we will see later, the VB approximation on the joint distributions lead to marginal distributions of the phase from the exponential family, hence the use of conjugate priors remains important. Von-Mises distribution [12], [8] is a useful distribution over the unit circle defined from  $[-\pi, \pi]$  on the real line, which serves as a conjugate prior to the phase for our conditional data model. It has one complex parameter, which serves as both location and concentration parameter for the distribution and has been used in varied applications on directional data.

$$f(\phi|\kappa_0) \equiv \mathcal{M}(\kappa_0) = \frac{1}{2\pi I_0(|\kappa_0|)} e^{\Re\{\kappa_0 e^{-j\phi}\}} \quad (4)$$

where  $\kappa_0 \in \mathbb{C}$ . In the signal processing literature this distribution is also known as the Tikhonov distribution [12] and has been previously used to model the phase error from the estimated phase using a phase-locked loop (PLL). In this work Von-Mises/Tikhonov distribution is considered to be the prior distribution of the phase in a Bayesian framework which was explicitly introduced in [10]. As the Von-Mises/Tikhonov distribution is confined to the unit circle, means and variances outside  $[-\pi, \pi]$  are undefined. Hence circular moments are used which are defined as:

$$\mathbb{E}[\phi] = \angle \kappa_0 \quad (5)$$

$$\begin{aligned} Var_{cir}[\phi] &= 1 - \mathbb{E}[\cos(\phi - \mathbb{E}[\phi])] \\ &= 1 - \frac{I_1(|\kappa_0|)}{I_0(|\kappa_0|)} \end{aligned} \quad (6)$$

Limiting behaviour of this distribution under various conditions of its hyperparameter are as follows

$$\lim_{|\kappa_0| \rightarrow \infty} f(\phi|\kappa_0) = \mathcal{N}_\phi\left(\angle \kappa_0, \frac{1}{|\kappa_0|}\right) \quad (7)$$

$$\lim_{|\kappa_0| \rightarrow 0} f(\phi|\kappa_0) = \mathcal{U}_\phi([-\pi, \pi]) \quad (8)$$

where  $\mathcal{N}$  is the normal distribution and  $\mathcal{U}$  is the uniform distribution.

### III SINGLE SYMBOL TRANSMISSION

Now we present the inference procedure for the single symbol transmission, where exact inference was shown to be possible in [10]. This analysis forms the basis for our generalization into the multi-symbol cases, wherein the Variational Bayes Method is applied. The data generation model in this case is taken to be

$$f(x, l, \phi | g, \kappa_0, \alpha, \beta_0, a) = f(x | a, \phi, r, l) f(l | \alpha) f(\phi | \kappa_0) \quad (9)$$

where

$$f(\phi | \kappa_0) \equiv \mathcal{M}(\kappa_0) \quad (10)$$

$$f(l | \alpha) \equiv \mathcal{M}_{u_l}(1, \alpha) = l' \alpha \quad (11)$$

$$f(\mathbf{x} | a, \phi, r, l) \equiv \mathcal{CN}_{\mathbf{x}}(l' a g e^{j\phi}, r) \quad (12)$$

The marginal posterior distributions for phase

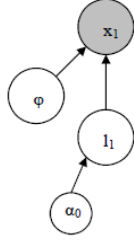


Fig. 1: **Graphical Model** - Single-symbol model in the case of uncertain phase

and the symbol can be easily calculated in this case and are given in [10] to be

$$f(\phi | \mathbf{x}) = \sum_l p_l \mathcal{M}(\kappa_l) \quad (13)$$

$$f(l | \mathbf{x}) = \mathcal{M}_{u_l}(1, p_l) \quad (14)$$

where

$$\kappa_l = \kappa_0 + \frac{2}{r} (l' a g)^H \mathbf{x} \quad (15)$$

$$= \kappa_0 + \frac{2}{r} (l' a)^H \sum_{k=1}^n s_k x_k e^{-j\omega k} \quad (16)$$

$$p_l \propto (l' \alpha) e^{(-\frac{1}{r} (l' a g)^H (l' a g))} I_0(|\kappa_l|) \quad (17)$$

The mean value of  $\phi$  for each component is given by  $\angle \kappa_l$ . Also  $\kappa_l$  can be characterized as a Discrete Time Fourier Transform on the incoming data, which is intuitive since  $\kappa_l$  is a sufficient statistic to the phase  $\phi$ . Here however we get more information than from a simple DTFT, namely about the confidence over the expected phase.

Another important observation about  $\kappa_l$  can be

seen with the following algebraic expansion:

$$|\kappa_l| = [\kappa_l \overline{\kappa_l}]^{1/2} \quad (18)$$

$$\begin{aligned} &= [\frac{4}{r^2} (l' a g)^H \mathbf{x} \overline{(l' a g)^H \mathbf{x}} \\ &\quad + \kappa_0 \overline{\kappa_0} + \frac{2}{r} \kappa_0 \overline{(l' a g)^H \mathbf{x}} + \frac{2}{r} \overline{\kappa_0} (l' a g)^H \mathbf{x}]^{1/2} \end{aligned} \quad (19)$$

From the above equation we can see that when  $\kappa_0 = 0$  (Uniform Prior), all except the first term vanishes. That implies then that when the prior is a uniform distribution, the decoding distribution is invariant to the symbols on the same radius in a QAM constellation, which is the origin of phase ambiguity. However a non-zero  $\kappa_0$  makes  $\kappa_l$  depend on the value of the symbols and hence results in a unique MAP from the decoding distribution.

### IV MULTI-SYMBOL TRANSMISSION

Real world communication systems involve multiple symbol periods, which are observed sequentially and a decoding distribution for the symbols is the desired output. The joint data model under the independent symbol sequence assumption (reasonable in turbo processing where pseudorandom inter-leavers are typically employed [9]) is then

$$f(\mathbf{x}_K | \phi, r, \alpha) = \prod_{i=1}^K \sum_{l_i} f(l_i | \alpha) f(x_i | a, \phi, r, l_i) \quad (20)$$

As a consequence of this product over a summation,  $M^K$  term mixture model appears after  $K$  time periods, a combinatorial explosion which leads to intractable exact inference.

#### a) Variational Bayes Method

A deterministic approximation, known as the Variational Bayes Method is employed to deal with this computational intractability. The Method aims at minimizing the Kullback-Leibler Divergence between the candidate approximation and the original joint distribution. This divergence can be interpreted as an information difference [15] between the two distributions. The intuition then is to get a functional form of the posterior marginals, so as to minimize the information lost due to the forced independence [13]. These methods have found many applications in machine learning, artificial intelligence and more recently in signal processing [9]. The essential result which forms the basis of this method is as follows:

**Theorem IV.1** *Variational Bayes Method [14]: Let  $f(\theta | \mathbf{x})$  be the posterior distribution of multi-variate parameter,  $\theta$ . The latter is partitioned into  $q$  sub-vectors of parameters or  $q$ -nodes:*

$$\theta = [\theta'_1, \theta'_2, \dots, \theta'_q]' \quad (21)$$

Let  $\check{f}(\theta|\mathbf{x})$  be an approximate distribution restricted to the set of conditionally independent distributions for  $\theta_1, \theta_2, \dots, \theta_q$ :

$$\check{f}(\theta|\mathbf{x}) = \prod_{i=1}^q \check{f}(\theta_i|\mathbf{x}) \quad (22)$$

Then, the minimum of Kullback-Leibler divergence  $\mathbf{KL}(\check{f}(\theta|\mathbf{x})||f(\theta|\mathbf{x}))$  is reached for

$$\check{f}(\theta_i|\mathbf{x}) \propto e^{\left(\mathbb{E}_{\check{f}(\theta_i|\mathbf{x})}[\ln(f(\theta, \mathbf{x}))]\right)} \quad (23)$$

where  $i = 1, \dots, q$  and  $\theta_{/i}$  in  $\theta$ , and  $\check{f}(\theta_{/i}|\mathbf{x}) = \prod_{j=1, j \neq i}^q \check{f}(\theta_j|\mathbf{x})$ . Now  $\check{f}(\theta|\mathbf{x})$  is the VB-approximation, and  $\check{f}(\theta_i|\mathbf{x})$  are the VB-marginals.

#### b) Offline VB Approximation

Consider the situation, where we observe a batch of symbol transfer periods together and are required to infer the marginal distributions for all the transferred symbols and the phase parameter. Prior independence of the symbols is assumed. However it should be noted, that a posteriori the symbols are independent conditional on  $\phi$ . Expanding the joint probability, over all unknown parameters, we have

$$\begin{aligned} f(\mathbf{x}_K, \mathbf{l}_K, \phi|\kappa_0, \alpha, r) \\ = f(\phi|\kappa_0) \prod_{i=1}^K f(x_i|a, \phi, r, l_i) f(l_i|\alpha) \end{aligned} \quad (24)$$

where

$$f(\phi|\kappa_0) \equiv \mathcal{M}(\kappa_0) \quad (25)$$

$$f(l_i|\alpha) \equiv \mathcal{Mu}_{l_i}(1, \alpha) = l'_i \alpha \quad (26)$$

$$f(x_i|a, \phi, r, l_i) \equiv \mathcal{CN}_{x_i}(l'_i a g_i e^{j\phi}, r) \quad (27)$$

From the above joint distribution, we find the

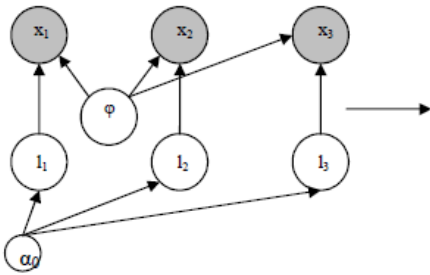


Fig. 2: **Graphical Model** - Multi-symbol model in the case of uncertain phase

posterior distributions for the label field and the phase. Integrating out the label field entails a  $M^K$  summation, making it intractable [13]. We therefore resort to using Theorem II.1 and derive the

approximate posterior distributions of the parameters as

$$\begin{aligned} \tilde{f}(\phi|\mathbf{x}_K) &\propto e^{\left(\mathbb{E}_{\tilde{f}(\phi|\mathbf{x}_K)}[\ln f(\mathbf{x}_K, \mathbf{l}_K, \phi|\kappa_0, \alpha, r)]\right)} \\ &= \mathcal{M}_\phi(\kappa_{\mathbf{l}_K}) \end{aligned} \quad (28)$$

$$\begin{aligned} \tilde{f}(l_i|\mathbf{x}_K) &\propto e^{\left(\mathbb{E}_{\tilde{f}(\phi|\mathbf{x}_K)}[\ln f(\mathbf{x}_K, \mathbf{l}_K, \phi|\kappa_0, \alpha, r)]\right)} \\ &= \mathcal{Mu}_{l_i}(1, p_{l_i}) \end{aligned} \quad (29)$$

with the shaping parameters as

$$\kappa_{\mathbf{l}_K} = \kappa_0 + \frac{2}{r} \sum_{i=1}^K (\hat{l}'_i a g_i)^H x_i \quad (30)$$

$$p_{l_i} \propto (l'_i \alpha) e^{\left(-\frac{1}{r} (l'_i a g_i)^H (l'_i a g_i) + \frac{2}{r} \delta_i\right)} \quad (31)$$

and the associated VB moment as

$$\begin{aligned} \delta_i &= \mathbb{E}_{\tilde{f}(\phi|\mathbf{x}_K)} [\Re \{ (l'_i a g_i)^H x_i e^{-j\phi} \}] \\ &= \Re \left\{ (l'_i a g_i)^H x_i e^{-j\phi_{\kappa_{\mathbf{l}_K}}} \right\} \frac{I_1(|\kappa_{\mathbf{l}_K}|)}{I_0(|\kappa_{\mathbf{l}_K}|)} \end{aligned} \quad (32)$$

$\kappa_{\mathbf{l}_K}$  again can be characterized as a DTFT. Important subsets of this method discussed in the literature include [9] where the authors have assumed a uniform prior to the phase which implies  $\kappa_0 = 0$ . Further if we make the ratio

$$\frac{I_1(|\kappa_{\mathbf{l}_K}|)}{I_0(|\kappa_{\mathbf{l}_K}|)} = 1$$

we get the EM-algorithm based turbo-synchronization routine as discussed in [3].

#### c) Online VB Approximation

In this case we observe each symbol transfer period sequentially and are required to infer the marginal distributions for each transferred symbol and phase parameter, after each symbol period. Again prior independence of the symbols is assumed. To deal with the increasing combinatorial complexity after each time step we conflate the posterior distribution of  $\phi$  which is a mixture of  $M$  Von-Mises components using Variational Bayes Method in to a single Von-Mises posterior. The associated sufficient statistic can then be used as the prior for the next symbol period. Expanding the joint probability, over unknown parameters after each time step, we have

$$f(x_t, l_t, \phi|\mathbf{x}_{t-1}) = f(x_t|a, \phi, r, l_t) f(l_t|\alpha) \tilde{f}(\phi|\mathbf{x}_{t-1}) \quad (33)$$

where

$$f(\phi|\mathbf{x}_0) = f(\phi|\kappa_0) \equiv \mathcal{M}(\kappa_0) \quad (34)$$

$$f(l_t|\alpha) \equiv \mathcal{Mu}_{l_t}(1, \alpha) = l'_t \alpha \quad (35)$$

$$f(x_t|a, \phi, r, l_t) \equiv \mathcal{CN}_{x_t}(l'_t a g_t e^{j\phi}, r) \quad (36)$$

On equation (13) from Section III of the posterior distribution of the phase, we apply the VB Method and derive the approximate conflated posterior distribution as

$$\begin{aligned}\tilde{f}(\phi|\mathbf{x}_{t-1}) &\propto e^{(\mathbb{E}_{\tilde{f}(l_{t-1}|x_{t-1})}[\ln f(x_{t-1}, l_{t-1}, \phi|\mathbf{x}_{t-2}, \kappa_0, \alpha, r)])} \\ &= \mathcal{M}_\phi(\kappa_{l_{t-1}})\end{aligned}\quad (37)$$

$$(38)$$

with the shaping parameter as

$$\hat{\kappa}_{l_{t-1}} = \hat{\kappa}_{l_{t-2}} + \frac{2}{r}(\hat{l}'_{t-1}ag_{t-1})^H x_{t-1} \quad (39)$$

$$p_{l_{t-1}} \propto (l'_{t-1}\alpha)e^{(-\frac{1}{r}(l'_{t-1}ag)^H(l'_{t-1}ag))}I_0(|\kappa_{l_{t-1}}|) \quad (40)$$

$$\hat{l}_{t-1} = \mathbb{E}[l_{t-1}] \quad (41)$$

Using this shaping parameter as initialization for the following symbol period we get from the single symbol transmission case:

$$f(\phi|\mathbf{x}_t) = \sum_l p_{l_t} \mathcal{M}(\kappa_{l_t}) \quad (42)$$

$$f(l_t|\mathbf{x}_t) = \mathcal{M}u_{l_t}(1, p_{l_t}) \quad (43)$$

where

$$\kappa_{l_t} = \hat{\kappa}_{l_{t-1}} + \frac{2}{r}(l'_t ag_t)^H x_t \quad (44)$$

$$p_{l_t} \propto (l'_t \alpha)e^{(-\frac{1}{r}(l'_t ag)^H(l'_t ag))}I_0(|\kappa_{l_t}|) \quad (45)$$

$\kappa_{l_t}$  again can be characterized as a DTFT as earlier.

## V SIMULATION

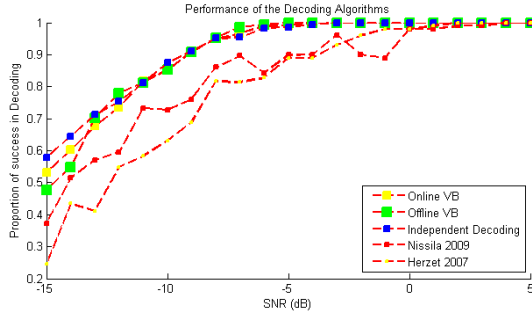


Fig. 3: - Performance of the VB-based decoding algorithms in the multi-symbol case

The above figure describes the proportion of successful identifications of the symbol transferred for various values of SNR, using the methods developed in this paper. In each batch 20 symbols are transmitted through an AWGN channel. In the low-SNR regime, from -15dB to 5dB we see that the Bayesian algorithms are significantly better. Independent decoding algorithm uses the exact decoding algorithm over each subsequent symbol period ignoring the previous periods. This method

turns out to be the fastest and most accurate for this simulation. The Online VB and the Offline VB approximations provide almost as good results, though they would be more robust in a more general setting. We can also clearly see the performance of [9] and [3] is significantly lower with respect to the decoding success. It should be noted, to deal with the problem of phase ambiguity for these two methods, 5 pilot symbols are used in front of the original 20.

## VI DISCUSSION

We get two main results from the analysis performed in this paper. Primarily we are able to extend the EM-Algorithm based synchronization into a fully Bayesian synchronization procedure. This can form the basis for future generalizations and analysis of even more complicated scenarios. Secondly, as a corollary to our analysis we are able to resolve the problem of phase ambiguity.

As shown in the Section III, setting  $\kappa_0 = 0$  renders  $\kappa_l$  independent of the expected phase  $\mathbb{E}[\phi]$ , which thereby results in the apparent ambiguity in the decoding distribution. The notion of non-zero  $\kappa_0$  seems quite inconsequential, but a concentrated prior however informs the presence of a unimodal phase distribution and regularizes the observation model which is intrinsically rotationally invariant for QAM. It must be noted however, that as the number of incoming observations increases the sampling distribution of the MAP symbol becomes relatively invariant to the value of  $\kappa_0$ , as long as  $\kappa_0 > 0$ .

Computationally it presents no overhead, as it involves only the addition of a constant. In contrast past approaches like Differential Coding [2], angle differential QAM scheme [6] and the use of pilot symbols [3] [4], among others which have been proposed to deal with this rotational invariance are visibly more complex and lack the simplicity of this observation.

In [3] the authors use the EM-Algorithm to deal with phase synchronization. This essentially implies using the sifting property of the Kronecker Delta function and thus using the expectations of the parameters in the associated functions rather than the expectations of the functions of the parameters itself. Significantly lower accuracy of the decoding distribution in the low-SNR regime is observed which converges in performance with increasing SNR.

## VII CONCLUSIONS

In this paper we have been able to extend the EM framework for tackling the phase-synchronization problem, to a fully Bayesian VB based algorithm wherein as a virtue of its modelling ability the problem of phase ambiguity does not arise. This

makes the concept of using "soft" information truly complete as no statistical information is lost at any point in the iterations. Further the basic framework developed motivates further relaxations into more general channels, such as those with multi-path fading, like Rician Fading [2]. In future work we would also be relaxing the condition on known noise parameter and deal with it in a Bayesian framework.

### VIII ACKNOWLEDGMENTS

This research was supported by Science Foundation Ireland, grant 08/RFP/MTH1710

### REFERENCES

- [1] PERSI DIACONIS AND DONALD YLVISAKER. **Conjugate Priors for Exponential Families**. *The Annals of Statistics*, **7**(2):pp. 269–281, 1979.
- [2] ANDREA GOLDSMITH. *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2005.
- [3] C. HERZET, N. NOELS, V. LOTTICI, H. WYMEERSCH, M. LUISE, M. MOENECLAËY, AND L. VANDENDORPE. **Code-Aided Turbo Synchronization**. *Proceedings of the IEEE*, **95**(6):1255–1271, june 2007.
- [4] C. HERZET, K. WORADIT, H. WYMEERSCH, AND L. VANDENDORPE. **Code-Aided Maximum-Likelihood Ambiguity Resolution Through Free-Energy Minimization**. *Signal Processing, IEEE Transactions on*, **58**(12):6238–6250, Dec 2010.
- [5] CHRISTIAN HIPPI. **Sufficient Statistics and Exponential Families**. *The Annals of Statistics*, **2**(6):pp. 1283–1292, 1974.
- [6] JENG-KUANG HWANG, YU-LUN CHIU, AND CHIA-SHU LIAO. **Angle differential-QAM scheme for resolving phase ambiguity in continuous transmission system**. *Int. J. Commun. Syst.*, **21**:631–641, June 2008.
- [7] MICHAEL I. JORDAN. **An introduction to variational methods for graphical models**. In *Machine Learning*, pages 183–233. MIT Press, 1999.
- [8] C. G. KHATRI AND K. V. MARDIA. **The Von Mises-Fisher Matrix Distribution in Orientation Statistics**. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**(1):pp. 95–106, 1977.
- [9] MAURI NISSILA AND SUBBARAYAN PASUPATHY. **Adaptive iterative detectors for phase-uncertain channels via variational bounding**. *Trans. Comm.*, **57**:716–725, March 2009.
- [10] ANTHONY QUINN, JEAN-PIERRE BARBOT, AND PASCAL LARZABAL. **The Bayesian Inference of Phase**. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2011, ICASSP'11*, 2011.
- [11] C.E. SHANNON. **Communication in the presence of noise**. *Proceedings of the IEEE*, **72**(9):1192 – 1201, sept 1984.
- [12] YURIY S. SHMALIY. **Von Mises/Tikhonov-based distributions for systems with differential phase measurement**. *Signal Processing*, **85**(4):693 – 703, 2005.
- [13] V. SMIDL AND A. QUINN. **The Variational Bayes Approximation In Bayesian Filtering**. In *IEEE International Conference on Acoustics, Speech and Signal Processing, 2006*, **3**, page III, May 2006.
- [14] VACLAV SMIDL AND ANTHONY QUINN. *The Variational Bayes Method in Signal Processing*. Springer, Dublin, Ireland, 2006.
- [15] MARTIN J. WAINWRIGHT AND MICHAEL I. JORDAN. **Graphical Models, Exponential Families, and Variational Inference**. *Foundations and Trends in Machine Learning*, **1**:1–305, 2008.